

# A Parametric Model for Doubly Interval Censored Lifetime

Yue Fang Loh<sup>1\*</sup>, Jayanthi Arasan<sup>1,2</sup>, Habshah Midi<sup>1,2</sup> and Mohd Rizam Abu Bakar<sup>1,2</sup>

<sup>1\*</sup>Department of Mathematics, Faculty of Science, Universiti Putra Malaysia, Malaysia.

<sup>2</sup>Institute for Mathematical Research, Universiti Putra Malaysia, Malaysia.

Received 14 August 2017; accepted 28 November 2017; available online 30 November 2017

**Abstract:** Doubly interval censored data is defined as elapsed time between two related events that is subject to interval or right censoring. In this paper, we extended a parametric model to incorporates doubly interval-, interval-, right censored and uncensored lifetime data. We assumed the initial event time follows uniform distribution and the lifetime follows the log logistic distribution. The interval censored event times are imputed using midpoint of their intervals for ease of the estimation process. The estimation procedure is studied at different sample sizes and attendance probabilities using simulated data. Finally, we study the Wald method of constructing confidence interval estimates for the parameters of the model. Conclusions were drawn based on the coverage probability study.

**Keyword:** doubly interval censored; log logistic; maximum likelihood estimation; midpoint imputation; Wald interval.

## 1. Introduction

Doubly interval censored (DIC) data mostly arise in epidemiology study due to the natural of disease or the structure of the study design. Let  $T$  be the lifetime of interest, if  $V$  is the initial event time and  $W$  is the subsequent event time, then  $V \in (V_L, V_R]$ ,  $W \in (W_L, W_R]$  and  $T = W - V$ , with  $V_L \leq V_R$  and  $W_L \leq W_R$ . A special case of DIC data occurs when  $V$  is interval censored (IC) and  $W$  is right censored (RC). In this paper, we named the lifetime as doubly interval censored-2 (DIC<sub>2</sub>) to differentiate it from DIC data. The doubly interval censored data also includes usual interval censored, right censored and uncensored (UC) data as special cases.

The aim of this study is to extend the log logistic model to incorporate DIC data and its special cases. The log logistic distribution is useful to model lifetime in survival analysis due to its ability to accommodate nonmonotonic hazard function. It is usually used to analyze lifetime of cancer patients, see [1,2].

The analysis of doubly interval censored data begins when De Gruttola and Lagakos [3] proposed a nonparametric estimation procedure based on the Turnbull's self-consistency algorithm. Following that, the analysis of doubly interval censored data has been studied extensively using nonparametric

and semiparametric regression approaches. Reich et al. [4] proposed that the likelihood contribution for a doubly interval censored lifetime is

$$\int_{v_L}^{v_R} \int_{w_L}^{w_R} f_V(v) f_T(w - v) dw dv,$$

where  $f_V(v)$  and  $f_T(t)$  are density function of  $V$  and  $T$  respectively. They assumed the initial event time follows uniform distribution and the lifetime of interest follows log normal distribution.

Kiani and Arasan [5] adapted Reich et al.'s idea and proposed a parametric model by assuming both initial event time and lifetime follow exponential distribution. In this paper, we follow Kiani and Arasan's procedure and assumed that the initial event time follows uniform distribution and lifetime follows log logistic distribution.

Researchers often apply imputation on the doubly interval censored data in order to ease the estimation process. For instance, midpoint imputation on interval censored initial event times in [6,7]. Law and Brookmeyer [8] pointed out that midpoint imputation is a reasonably adequate procedure for interval widths of 2 years or less if the median of lifetime of interest is 10 years.

\*Corresponding author: [yuefangloh@yahoo.com](mailto:yuefangloh@yahoo.com)

2017 UTHM Publisher. All right reserved.

[penerbit.uthm.edu.my/ojs/index.php/jst](http://penerbit.uthm.edu.my/ojs/index.php/jst)

## 2. The model

We assume the initial event,  $V \sim U(a, b)$  and the lifetime,  $T \sim LL(\lambda, \gamma)$  where  $\lambda$  is scale parameter and  $\gamma$  is shape parameter. The density and survival functions of  $V$  are given by

$$f_V(v) = \frac{1}{b-a},$$

$$S_V(v) = \frac{b-v}{b-a}, v > 0.$$

The density and survival functions of  $T$  are given by

$$f_T(t) = \frac{e^{\lambda\gamma} t^{\gamma-1}}{(1 + e^{\lambda\gamma} t^{\gamma})^2}, \lambda \in \mathbb{R}, \gamma > 0, t > 0,$$

$$S_T(t) = \frac{1}{1 + e^{\lambda\gamma} t^{\gamma}}, \lambda \in \mathbb{R}, \gamma > 0, t > 0.$$

For the case where both  $V$  and  $W$  are interval censored,  $T$  is doubly interval censored. We impute both  $V$  and  $W$  in order to reduce  $T$  to uncensored data with  $t_{UC'_i} = w'_i - v'_i$  for  $i = 1, 2, \dots, n$  where  $v'_i = (v_{L_i} + v_{R_i})/2$  and  $w'_i = (w_{L_i} + w_{R_i})/2$ . The likelihood contribution would then be  $f_T(t_{UC'_i})$ . For the case where  $V$  is interval censored and  $W$  is right censored,  $T$  is doubly interval censored-2, we impute  $V$  to reduce  $T$  to right censored data with  $t_{RC'_i} = w_{L_i} - v'_i$ . The likelihood contribution would then be  $S_T(t_{RC'_i})$ .

If either  $V$  or  $W$  is uncensored while the other is interval censored,  $T$  becomes interval censored. The interval  $(t_{L_i}, t_{R_i}]$  is equal to  $(w_i - v_{R_i}, w_i - v_{L_i}]$  when  $V$  is interval censored; and  $(w_{L_i} - v_i, w_{R_i} - v_i]$  when  $W$  is interval censored. For this case, the likelihood contribution is  $\int_{t_{L_i}}^{t_{R_i}} f_T(t) dt = S_T(t_{L_i}) - S_T(t_{R_i})$ . If  $V$  is uncensored and  $W$  is right censored,  $T$  becomes right censored and  $t_{RC_i} = w_{L_i} - v_i$ . For this case, the likelihood contribution would be  $S_T(t_{RC_i})$ . If both  $V$  and  $W$  are uncensored,  $T$  becomes uncensored and  $t_i = w_i - v_i$ . For this case, the likelihood contribution is  $f_T(t_i)$ .

Let us define the following censoring indicator variables,

$$\begin{aligned} \delta_{1i} &= 1 \text{ if } T \text{ is DIC, } 0 \text{ otherwise;} \\ \delta_{2i} &= 1 \text{ if } T \text{ is DIC}_2, 0 \text{ otherwise;} \\ \delta_{3i} &= 1 \text{ if } T \text{ is IC, } 0 \text{ otherwise;} \\ \delta_{4i} &= 1 \text{ if } T \text{ is RC, } 0 \text{ otherwise;} \\ \delta_{5i} &= 1 \text{ if } T \text{ is UC, } 0 \text{ otherwise.} \end{aligned}$$

Note that  $\delta_{5i} = 1 - (\delta_{1i} + \delta_{2i} + \delta_{3i} + \delta_{4i})$ . Then, the likelihood function for the full sample is written as

$$L(\lambda, \gamma) = \prod_{i=1}^n \left[ f_T(t_{UC'_i}) \right]^{\delta_{1i}} \times \left[ S_T(t_{RC'_i}) \right]^{\delta_{2i}} \times \left[ S_T(t_{L_i}) - S_T(t_{R_i}) \right]^{\delta_{3i}} \times \left[ S_T(t_{RC_i}) \right]^{\delta_{4i}} \times \left[ f_T(t_{UC_i}) \right]^{\delta_{5i}}.$$

The first and second partial derivatives of log likelihood function  $\ell(\lambda, \gamma)$  is given as follows,

$$\begin{aligned} \frac{\partial \ell(\lambda, \gamma)}{\partial \lambda} &= \sum_{i=1}^n \left\{ \delta_{1i} \left[ 1 - \frac{2e^{\lambda\gamma} t_{UC'_i}^{\gamma}}{1 + e^{\lambda\gamma} t_{UC'_i}^{\gamma}} \right] - \right. \\ &\delta_{2i} \left[ \frac{e^{\lambda\gamma} t_{RC'_i}^{\gamma}}{1 + e^{\lambda\gamma} t_{RC'_i}^{\gamma}} \right] + \delta_{3i} \left[ 1 - \frac{e^{\lambda\gamma} t_{L_i}^{\gamma}}{1 + e^{\lambda\gamma} t_{L_i}^{\gamma}} - \frac{e^{\lambda\gamma} t_{R_i}^{\gamma}}{1 + e^{\lambda\gamma} t_{R_i}^{\gamma}} \right] - \\ &\left. \delta_{4i} \left[ \frac{e^{\lambda\gamma} t_{RC_i}^{\gamma}}{1 + e^{\lambda\gamma} t_{RC_i}^{\gamma}} \right] + \delta_{5i} \left[ \frac{2e^{\lambda\gamma} t_i^{\gamma}}{1 + e^{\lambda\gamma} t_i^{\gamma}} \right] \right\}, \end{aligned}$$

$$\begin{aligned} \frac{\partial \ell(\lambda, \gamma)}{\partial \gamma} &= \sum_{i=1}^n \left\{ \delta_{1i} \left[ \frac{1}{\gamma} + \ln t_{UC'_i} - \right. \right. \\ &\left. \frac{2e^{\lambda\gamma} t_{UC'_i}^{\gamma} \ln t_{UC'_i}}{1 + e^{\lambda\gamma} t_{UC'_i}^{\gamma}} \right] - \delta_{2i} \left[ \frac{e^{\lambda\gamma} t_{RC'_i}^{\gamma} \ln t_{RC'_i}}{1 + e^{\lambda\gamma} t_{RC'_i}^{\gamma}} \right] + \\ &\delta_{3i} \left[ \frac{t_{R_i}^{\gamma} \ln t_{R_i} - t_{L_i}^{\gamma} \ln t_{L_i}}{t_{R_i}^{\gamma} - t_{L_i}^{\gamma}} - \frac{e^{\lambda\gamma} t_{L_i}^{\gamma} \ln t_{L_i}}{1 + e^{\lambda\gamma} t_{L_i}^{\gamma}} - \right. \\ &\left. \frac{e^{\lambda\gamma} t_{R_i}^{\gamma} \ln t_{R_i}}{1 + e^{\lambda\gamma} t_{R_i}^{\gamma}} \right] - \delta_{4i} \left[ \frac{e^{\lambda\gamma} t_{RC_i}^{\gamma} \ln t_{RC_i}}{1 + e^{\lambda\gamma} t_{RC_i}^{\gamma}} \right] + \\ &\left. \delta_{5i} \left[ \frac{1}{\gamma} + \ln t_i - \frac{2e^{\lambda\gamma} t_i^{\gamma} \ln t_i}{1 + e^{\lambda\gamma} t_i^{\gamma}} \right] \right\}, \end{aligned}$$

$$\begin{aligned} \frac{\partial^2 \ell(\lambda, \gamma)}{\partial \lambda^2} &= \sum_{i=1}^n \left\{ 2\delta_{1i} \left[ \frac{e^{\lambda\gamma} t_{UC'_i}^{\gamma}}{(1 + e^{\lambda\gamma} t_{UC'_i}^{\gamma})^2} \right] + \right. \\ &\delta_{2i} \left[ \frac{e^{\lambda\gamma} t_{RC'_i}^{\gamma}}{(1 + e^{\lambda\gamma} t_{RC'_i}^{\gamma})^2} \right] + \delta_{3i} \left[ \frac{e^{\lambda\gamma} t_{L_i}^{\gamma}}{(1 + e^{\lambda\gamma} t_{L_i}^{\gamma})^2} + \right. \\ &\left. \frac{e^{\lambda\gamma} t_{R_i}^{\gamma}}{(1 + e^{\lambda\gamma} t_{R_i}^{\gamma})^2} \right] + \delta_{4i} \left[ \frac{e^{\lambda\gamma} t_{RC_i}^{\gamma}}{(1 + e^{\lambda\gamma} t_{RC_i}^{\gamma})^2} \right] + \\ &\left. \delta_{5i} \left[ \frac{2e^{\lambda\gamma} t_i^{\gamma}}{(1 + e^{\lambda\gamma} t_i^{\gamma})^2} \right] \right\} \end{aligned}$$

$$\begin{aligned}
& \left. \frac{e^{\lambda t_{R_i}^\gamma}}{(1+e^{\lambda t_{R_i}^\gamma})^2} \right] + \delta_{4i} \left[ \frac{e^{\lambda t_{RC_i}^\gamma}}{(1+e^{\lambda t_{RC_i}^\gamma})^2} \right] + \\
& 2\delta_{5i} \left[ \frac{e^{\lambda t_i^\gamma}}{(1+e^{\lambda t_i^\gamma})^2} \right] \Bigg\}, \\
\frac{\partial^2 \ell(\lambda, \gamma)}{\partial \lambda \partial \gamma} &= \sum_{i=1}^n - \left\{ 2\delta_{1i} \left[ \frac{e^{\lambda t_{UC_i}^\gamma} \ln t_{UC_i}}{(1+e^{\lambda t_{UC_i}^\gamma})^2} \right] + \right. \\
& \delta_{2i} \left[ \frac{e^{\lambda t_{RC_i}^\gamma} \ln t_{RC_i}}{(1+e^{\lambda t_{RC_i}^\gamma})^2} \right] + \delta_{3i} \left[ \frac{e^{\lambda t_{L_i}^\gamma} \ln t_{L_i}}{(1+e^{\lambda t_{L_i}^\gamma})^2} \right] + \\
& \left. \frac{e^{\lambda t_{R_i}^\gamma} \ln t_{R_i}}{(1+e^{\lambda t_{R_i}^\gamma})^2} \right] + \delta_{4i} \left[ \frac{e^{\lambda t_{RC_i}^\gamma} \ln t_{RC_i}}{(1+e^{\lambda t_{RC_i}^\gamma})^2} \right] + \\
& 2\delta_{5i} \left[ \frac{e^{\lambda t_i^\gamma} \ln t_i}{(1+e^{\lambda t_i^\gamma})^2} \right] \Bigg\}, \\
\frac{\partial^2 \ell(\lambda, \gamma)}{\partial \gamma^2} &= \sum_{i=1}^n - \left\{ \delta_{1i} \left[ \frac{1}{\gamma^2} + \right. \right. \\
& \left. \frac{2e^{\lambda t_{UC_i}^\gamma} (\ln t_{UC_i})^2}{(1+e^{\lambda t_{UC_i}^\gamma})^2} \right] + \delta_{2i} \left[ \frac{e^{\lambda t_{RC_i}^\gamma} (\ln t_{RC_i})^2}{(1+e^{\lambda t_{RC_i}^\gamma})^2} \right] + \\
& \delta_{3i} \left[ \frac{t_{R_i}^\gamma t_{L_i}^\gamma (\ln t_{R_i} - \ln t_{L_i})^2}{(t_{R_i}^\gamma - t_{L_i}^\gamma)^2} + \frac{e^{\lambda t_{L_i}^\gamma} (\ln t_{L_i})^2}{(1+e^{\lambda t_{L_i}^\gamma})^2} + \right. \\
& \left. \frac{e^{\lambda t_{R_i}^\gamma} (\ln t_{R_i})^2}{(1+e^{\lambda t_{R_i}^\gamma})^2} \right] + \delta_{4i} \left[ \frac{e^{\lambda t_{RC_i}^\gamma} (\ln t_{RC_i})^2}{(1+e^{\lambda t_{RC_i}^\gamma})^2} \right] + \\
& \left. \delta_{5i} \left[ \frac{1}{\gamma^2} + \frac{2e^{\lambda t_i^\gamma} (\ln t_i)^2}{(1+e^{\lambda t_i^\gamma})^2} \right] \right\}.
\end{aligned}$$

The maximum likelihood estimator (MLE) of the parameters is obtained using Newton-Raphson algorithm.

### 3. Simulation study

A simulation study using  $N = 1500$  replications of sizes,  $n = 50, 250$  and  $450$  was conducted to examine the estimation procedure. The initial event time,  $V$  is assumed to follows  $U(0,16)$  and the lifetime,  $T$  is assumed to be log logistically distributed with parameters  $\lambda$  and  $\gamma$ . The values  $-4.3$  and  $2$

were chosen as the parameter value of  $\lambda$  and  $\gamma$  respectively. We assumed the study lasted for 60 months with examination scheduled on monthly basis. We further assumed that subjects will attend scheduled examination with attendance probabilities,  $q = 1.0$  and  $0.7$ .

For each subject in the sample,  $v_i$  is simulated from  $U(0,16)$ . A random number  $u_{1i}$  is generated from  $U(0,1)$  to produce  $t_i$  where

$$t_i = \left[ e^{-\lambda} \left( \frac{1}{u_{1i}} - 1 \right) \right]^{1/\gamma}.$$

Then  $w_i = t_i + v_i$  is calculated. In order to obtain the intervals  $(v_{L_i}, v_{R_i})$  and  $(w_{L_i}, w_{R_i})$  for  $v_i$  and  $w_i$  respectively, we generated a sequence of potential examination times and a sequence of actual examination times. Assuming all subjects will have the same sequence of potential examination times,  $PE = (pe_1, pe_2, \dots, pe_{60})$ . Subjects will attend examination at each  $pe_j$  with attendance probabilities  $q$  where  $j = 2, 3, \dots, 60$ . Hence each subject will have their own sequence of actual examination times,  $AE_i = (ae_{i1}, ae_{i2}, \dots, ae_{ih_i})$  where  $0 \leq h_i \leq 60$  which is simulated from the Bernoulli distribution with predefined attendance probabilities. Following this, we simulated random number  $u_j$  from  $U(0,1)$  where  $j = 2, 3, \dots, 60$  and assume  $u_1 = 0$ . In this simulation study, we assumed all subject will not miss the first scheduled examination time, hence  $ae_{i1} = pe_1$ . We defined an indicator variable for  $pe_j$ 's,

$$I_j = \begin{cases} 1, & \text{if subject attend } pe_j \ (u_j \leq q); \\ 0, & \text{if subject miss } pe_j \ (u_j > q). \end{cases}$$

Then,  $h_i = \sum_{j=1}^{60} I_j$ . The intervals for  $v_i$  and  $w_i$  is obtained from  $AE_i$  using the following guidelines:

$v_{L_i}$  = largest element of  $AE_i$  which is less than  $v_i$ ;  
 $v_{R_i}$  = smallest element of  $AE_i$  which is greater than  $v_i$ ;  
 $w_{L_i}$  = largest element of  $AE_i$  which is less than  $w_i$ ;  
 $w_{R_i}$  = smallest element of  $AE_i$  which is greater than  $w_i$ .

If  $w_i > ae_{ih_i}$ , then  $W$  is right censored with  $(ae_{ih_i}, \infty)$ . We further defined two time-windows in order to randomly select some subjects that are uncensored on  $V$  or  $W$ . The time-window for uncensored  $V$  is

$$[G_{1i}, G_{2i}] = [v_{Li} + (v_{Ri} - v_{Li})u_{2i} - \epsilon, v_{Li} + (v_{Ri} - v_{Li})u_{2i} + \epsilon],$$

and for uncensored  $W$  is

$$[G_{3i}, G_{4i}] = [w_{Li} + (w_{Ri} - w_{Li})u_{3i} - \epsilon, w_{Li} + (w_{Ri} - w_{Li})u_{3i} + \epsilon],$$

where  $\epsilon = 0.25$  and  $u_{2i}$  and  $u_{3i}$  are random numbers generated from  $U(0,1)$ . If both  $v_i$  and  $w_i$  fall in the same interval, these observations are discarded. Two new values of  $v_i$  and  $t_i$  are generated to calculate  $w_i$  and repeat the above process again. This simulation algorithm will yield five possible types of data,

- 1) Both  $V$  and  $W$  are IC, then  $T$  is DIC;
- 2)  $V$  is IC,  $W$  is RC, then  $T$  is DIC<sub>2</sub>;
- 3a)  $V$  is IC,  $W$  is UC, then  $T$  is IC;
- 3b)  $V$  is UC,  $W$  is IC, then  $T$  is IC;
- 4)  $V$  is UC,  $W$  is RC, then  $T$  is RC;
- 5) Both  $V$  and  $W$  are UC, then  $T$  is UC.

Before we proceed to the estimation, we imputed  $V$  and  $W$  with the midpoint of their interval when the lifetime,  $T$  is doubly interval censored or doubly interval censored-2.

### 3.1 Simulation results

Table 1 shows the proportion of different types of data generated with different attendance probabilities. From the results, we noticed that more doubly interval censored and doubly interval censored-2 lifetime data is observed at lower attendance probabilities. This is due to the fact that chances of producing interval censored data on both  $V$  and  $W$  are higher at wider intervals. Also, the average percentage of uncensored lifetime data increase with an increase in the attendance probabilities. This is due to the fact that chances of producing uncensored data on both  $V$  and  $W$  are higher at narrower intervals. Similar to interval censored lifetime data, the chances of producing uncensored data on either  $V$  or  $W$  are higher at narrower intervals.

Therefore, more interval censored lifetime data is observed at higher attendance probabilities.

Table 2 gives the bias, standard error (SE) and root mean square error (RMSE) of the parameter estimates at various sample sizes and attendance probabilities. From the results, we observed that both bias and standard error values are relatively low for all parameter estimates. The values of bias, standard error and RMSE decrease with an increase in the sample size and attendance probabilities. Therefore, based on the RMSE of both parameter estimates, we concluded that the procedure performs well in estimating the parameters of the model. In addition, the procedure performs better at higher attendance probabilities.

Table 1 Average percentages of different types of data generated from simulation study

Attendance probabilities	1.0	0.7
$T$ is DIC	30.76	45.15
$T$ is DIC <sub>2</sub>	1.56	1.96
$T$ is IC	47.84	42.15
$T$ is RC	1.20	0.91
$T$ is UC	18.64	9.82

Table 2 Bias, SE and RMSE of the parameter estimates

Para. est.		$\hat{\lambda}$		$\hat{\gamma}$	
Attend. prob.		1.0	0.7	1.0	0.7
Bias	50	-0.2157	-0.3359	0.0908	0.1345
	250	-0.1011	-0.2628	0.0401	0.0982
	450	-0.0796	-0.2455	0.0322	0.0871
SE	50	0.5847	0.6038	0.2517	0.2584
	250	0.2550	0.2653	0.1057	0.1106
	450	0.1860	0.1878	0.0783	0.0790
RMSE	50	0.6232	0.6909	0.2676	0.2913
	250	0.2744	0.3734	0.1130	0.1479
	450	0.2023	0.3091	0.0847	0.1176

### 4. Confidence interval estimates

Let  $\hat{\theta}$  be the vector of MLEs for the vector of parameters  $\theta = (\lambda, \gamma)$ . Under the mild regularity conditions,  $\hat{\theta}$  is asymptotically normally distributed with mean  $\theta$  and variance  $I(\theta)^{-1}$ , where  $I(\theta)$  is the Fisher information matrix evaluated at  $\theta$  and is estimated by the observed information matrix evaluated at the MLEs,  $i(\hat{\theta})$ . For  $\lambda$ , the estimate of  $\text{var}(\hat{\lambda})$  can be obtained from the first diagonal element of the inverse of  $i(\hat{\theta})$ . Then, the  $100(1 - \alpha)\%$  CI for  $\lambda$  could be expressed as

$$\left( \hat{\lambda} - z_{1-\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\lambda})}, \hat{\lambda} + z_{1-\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\lambda})} \right)$$

where  $z_{1-\alpha/2}$  is the  $1 - \alpha/2$  quantile of the standard normal distribution. The Wald confidence interval estimates of  $\gamma$  is obtained in the similar manner.

#### 4.1 Coverage probability study

We conducted a coverage probability study using  $N = 1500$  replications with sample sizes  $n = 50, 250$  and  $450$  to compare the performance of the Wald confidence interval estimates at nominal error probabilities,  $\alpha = 0.05$  and  $0.10$ . The values of  $-4.3$  and  $2$  were chosen as the parameters of  $\lambda$  and  $\gamma$ . Other settings are similar to what had previously discussed in the simulation study. Following that, we calculate the total error probabilities by adding the number of times in which the interval did not contain the true parameter value divided by total replications. The estimated total error probability is obtained by adding left and right error probability. For  $\lambda$ , the left error probability for the Wald intervals is

$$\text{left} = \# \left( \hat{\lambda} - z_{1-\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\lambda})} > \lambda \right) / 1500,$$

and the right error probability is

$$\text{right} = \# \left( \hat{\lambda} + z_{1-\alpha/2} \sqrt{\widehat{\text{var}}(\hat{\lambda})} < \lambda \right) / 1500.$$

The left and right error probabilities for  $\gamma$  are obtained in the similar manner. Following Doganaksoy and Schmee [9], the method is anticonservative (AC) if the total error probability is greater than  $\alpha + 2.58 \times \text{SE}(\hat{\alpha})$ . The method is conservative (C) if the total error probability is less than  $\alpha - 2.58 \times \text{SE}(\hat{\alpha})$ . The method is asymmetrical (AS) if the larger error probability is greater than 1.5 times the smaller error. The  $\text{SE}(\hat{\alpha})$  is defined as the standard error of estimated error probability and is approximately  $\text{SE}(\hat{\alpha}) = \sqrt{\alpha(1-\alpha)/N}$ .

#### 4.2 Results and Discussion

Table 3 gives the summary of the performance of Wald intervals. At both  $\alpha = 0.05$  and  $0.10$ , the Wald intervals does not produce any anticonservative and conservative intervals when the attendance probability is 1.0, regardless of the sample size. However, the intervals produce asymmetrical intervals for both parameters, at all sample sizes. When the attendance probability is 0.7, the Wald intervals does not produce any conservative intervals. It produces anticonservative intervals for both parameters when  $n > 50$ . In addition, intervals are highly asymmetrical.

Table 3 Summary of the interval estimates at  $\alpha = 0.05$  and  $0.10$

$\alpha$	Attend. prob.	AC	C	AS
0.05	1.0	0	0	6
	0.7	4	0	6
0.10	1.0	0	0	6
	0.7	4	0	6

Table 4 gives the estimated error probabilities of the Wald confidence interval estimates. We could see that the total error probabilities are close to the nominal error probabilities when attendance probability is 1.0. The total error probabilities are far from the nominal when attendance probability is 0.7 and  $n > 50$ . The intervals produced rather asymmetrical intervals. Fig. 1 illustrates the behavior of left and right error probabilities around  $\alpha/2$ . Ideally, we want the error probabilities to be close to the  $\alpha/2$ . However, we clearly observed a substantial deviation from the  $\alpha/2$  when attendance probability is 0.7 and  $n > 50$ .

#### 5. Conclusion

In this paper, the MLE for the parameters of the log logistic model in the presence of doubly interval-, interval-, right censored and uncensored data were obtained. A simulation result indicates that the bias, standard error and RMSE value decrease when the attendance probabilities and sample sizes increase. The performance of Wald confidence interval estimates for parameters of the model is studied. It works well at both nominal levels when the attendance probabilities in 1.0. This method is known to produce asymmetrical intervals [10]. Thus, the Wald intervals might not be reliable in making inferences to the

parameters of this model. Therefore, other confidence interval estimation methods such as likelihood ratio, jackknife and bootstrap methods for the parameters of the model could

be investigated in future study. The model could easily be extended to include covariate information.

Table 4 Estimated error probabilities at  $\alpha = 0.05$  and 0.10

Para.	$q$	$n$	$\alpha = 0.05$			$\alpha = 0.10$		
			Left	Right	Total	Left	Right	Total
$\lambda$	1.0	50	0.0113	0.0313	0.0427	0.0247	0.0687	0.0933
		250	0.0053	0.0480	0.0533	0.0207	0.0907	0.1113
		450	0.0080	0.0487	0.0567	0.0147	0.0913	0.1060
	0.7	50	0.0040	0.0447	0.0487	0.0120	0.0860	0.0980
		250	0.0007	0.1380	0.1387 <sup>AC</sup>	0.0033	0.2280	0.2313 <sup>AC</sup>
		450	0.0020	0.1993	0.2013 <sup>AC</sup>	0.0020	0.3100	0.3120 <sup>AC</sup>
$\gamma$	1.0	50	0.0347	0.0120	0.0467	0.0613	0.0280	0.0893
		250	0.0407	0.0080	0.0487	0.0800	0.0193	0.0993
		450	0.0427	0.0093	0.0520	0.0847	0.0213	0.1060
	0.7	50	0.0400	0.0080	0.0480	0.0813	0.0160	0.0973
		250	0.1120	0.0013	0.1133 <sup>AC</sup>	0.1947	0.0033	0.1980 <sup>AC</sup>
		450	0.1440	0.0007	0.1447 <sup>AC</sup>	0.2347	0.0033	0.2380 <sup>AC</sup>

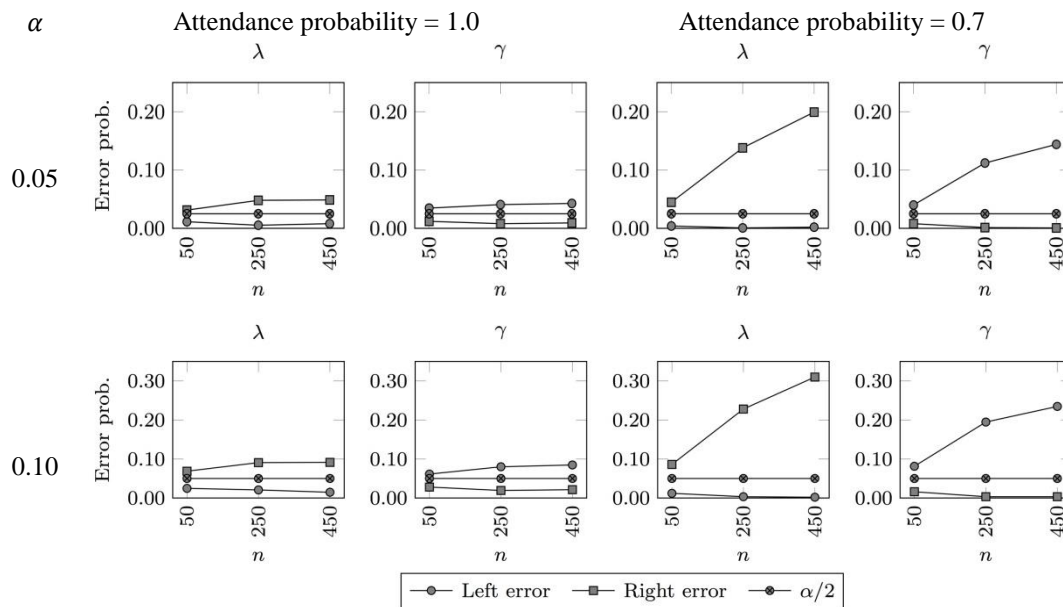


Fig. 1 Estimated error probabilities at  $\alpha = 0.05$  and 0.10

### Acknowledgement

We gratefully acknowledge financial support from the Ministry of Education Malaysia. The research leading to these results has received funding from the Fundamental Research Grant Scheme (FRGS), 5524673.

### References

- [1] Arasan, J. and Adam, M.B. (2014). "Double bootstrap confidence interval estimates with censored and truncated data" in Journal of Modern Applied Statistical Methods, Vol. 13. No. 2 pp. 399-419.
- [2] Gupta, R.C., Akman, O. and Lvin, S. (1999). "A study of log-logistic model in survival analysis" in Biometrical Journal, Vol. 41. No. 4 pp. 431-443.
- [3] De Gruttola, V. and Lagakos, S.W. (1989). "Analysis of doubly-censored survival data, with application to AIDS" in Biometrics, Vol. 45. No. 1 pp. 1-11.
- [4] Reich, N.G., Lessler, J., Cummings, D.A. and Brookmeyer, R. (2009). "Estimating incubation period distributions with coarse

- data” in *Statistics in medicine*, Vol. 28. No. 22 pp. 2769-2784.
- [5] Kiani, K. and Arasan, J. (2012). “Interval estimations for parameters of exponential model with doubly interval-censored survival time data” in *Proceedings of 2nd Regional Conference on Applied and Engineering Mathematics (RCAEM-II) 2012*, pp. 653–660. UniMap.
- [6] Lui, K.J., Darrow, W.W. and Rutherford, G. (1988). “A model-based estimate of the mean incubation period for AIDS in homosexual men” in *Science*, Vol. 240. No. 4857 pp. 1333-1335.
- [7] Mariotto, A.B., Mariotti, S., Pezzotti, P., Rezza, G. and Verdecchia, A. (1992). “Estimation of the acquired immunodeficiency syndrome incubation period in intravenous drug users: a comparison with male homosexuals” in *American Journal of Epidemiology*, Vol. 135. No. 4 pp. 428-437.
- [8] Law, C.G. and Brookmeyer, R. (1992). “Effects of mid-point imputation on the analysis of doubly censored data” in *Statistics in medicine*, Vol. 11. No. 12 pp. 1569-1578.
- [9] Doganaksoy, N. and Schmee, J. (1993). “Comparisons of approximate confidence intervals for distributions used in life-data analysis” in *Technometrics*, Vol. 35. No. 2 pp. 175-184.
- [10] Arasan, J. and Lunn, M. (2009). “Survival model of a parallel system with dependent failures and time varying covariates” in *Journal of Statistical Planning and Inference*, Vol. 139. No. 3 pp. 944-951.